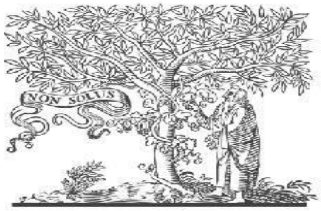
International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

COPY RIGHT



ELSEVIER SSRN

IJIEMR Transactions, online available on 10th Apr 2023. Link :http://www.ijiemr.org/downloads.php?vol=Volume-12&issue=Issue 04

10.48047/IJIEMR/V12/ISSUE 04/89

Title OBJECT DETECTION USING CNN

Volume 12, ISSUE 04, Pages: 728-734

Paper Authors

Mr.N.Ashok, Kasikala Charan Kumar, Kancharla Venkatesh , Mupparaju Edukondalu

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

OBJECT DETECTION USING CNN

Mr.N.Ashok[1], **assistant professor**, Department of IT,
Vasireddy Venkatadri Institute of Technology, Nambur, Guntur Dt., Andhra Pradesh.

Kasikala Charan Kumar[2], **Kancharla Venkatesh** [3], **Mupparaju Edukondalu**[4]
[2,3,4.] UG Students, Department of IT,
Vasireddy Venkatadri Institute of Technology, Nambur, Guntur Dt., Andhra Pradesh.
[1] charan.kasikala@gmail.com, [2]kancharlavenky07@gmail.com
[3]edukondalumupparaju1864@gmail.com

Abstract

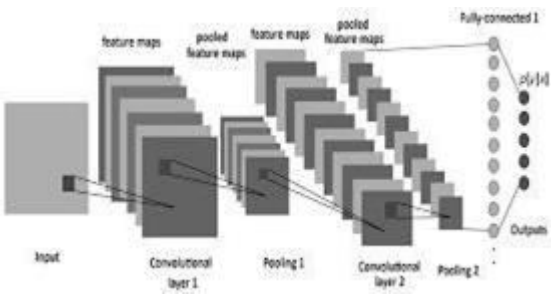
Detecting the existence, position, and posture of objects in an image or video is the task of object detection, a branch of computer vision. Real Time Objects. It is a crucial task in many applications, such as robotics, self-driving cars, and surveillance systems. Object detection methods in machine learning often utilize a mix of feature extraction and classification approaches to identify things. The process starts by extracting relevant features from the input image or video, such as edges, corners, and textures. These features are then fed into a classifier, which uses them to predict the class label of the object (e.g., car, pedestrian, stop sign). Convolutional neural networks (CNNs) are often used to analyse images and find areas that contain objects as a standard method for object detection. The CNN may be trained on a sizable dataset of annotated pictures and then used to anticipate the presence and placement of objects in fresh photos and videos.

Keywords: Convolutional neural network (CNN), Real time objects, annotated images, Object Detection, Computer Vision.

Introduction

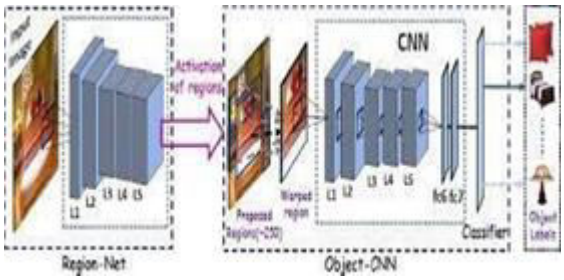
In computer vision systems, object detection is critical. It may be utilised for a variety of purposes, such as video surveillance [1]. CNN has made major advancements recently and is a sparkling pearl in the blooming deep neural network treasure house. Moreover, computer vision technologies have made it possible for artificial intelligence to analyse and perceive the visual aspects of the world. Deep learning-based computer vision algorithms have had great success in the past few years such as picture segmentation, object identification, and image classification are computer vision tasks. This accomplishment can be ascribed to improvements in computer hardware performance and the creation of massive picture annotation data sets. Nowadays, object detection is heavily employed in both academics and the real world, such as in the identification of video fires [2]. Conventional convolution entails applying the convolution kernel to the picture and computing the grey value of each pixel after a sequence of matrix operations.

Contrary to traditional convolution, transposed convolution is the process of convolution from low- dimensional feature mapping to high- dimensional feature mapping. It is frequently employed in semantic analysis [3]. Active learning is an additional approach for object detection [4]. Object identification, which involves both object recognition and picture categorization, is one of the interesting issues that computer vision poses. Convolutional neural networks, deep learning, and the increased parallelism processing power provided by graphics processing units have all made significant gains in the field of science in recent years, which has led to a significant body of research on these kinds of problems (GPUs) [5]. The most effective object categorization and detection algorithms nowadays employ deep learning architectures with several specialised layers for automating the filtering and feature extraction process[6]. Convolutional neural networks are heavily utilised, which increases the ability to learn the feature extraction phase observed in deep learning-based methods [7].



There are one or more convolutional layers in each convolutional neural network, which are computer programmes that extract useful information from the input image. Also, each convolution layer is composed of a number of filters, or square matrices, that traverse across the picture and record the weighted sum of the pixel values at different places. Every filter has a variety of settings and extracts different information from the input image. A convolution layer generates a set of "feature maps" [8]. Self-driving cars have dominated discussions about computer vision, a knowledge base topic that has gained a lot of momentum in recent years.[9]. Real-time object detection is an essential component of computer vision as well. Real-time object detection is helpful for a variety of tasks, including cause estimation, vehicle identification, traffic management, CCTV surveillance, etc. [10]

The capacity to provide vision on a computer is therefore made feasible by the necessity of daily technological advancement [11]. The power of the GPU is enormously needed for computer-based systems, and despite this, they aren't necessarily real-time, making them unsuitable for use in everyday applications. You Only Look Once (YOLO), faster R-CNN, faster R-CNN, and similar techniques [12].



Ultimately, the output of the convolution layers is flattened into a single dimension matrix that represents the numerical representation of the image's characteristics. This matrix is then fed into a succession of "fully connected" layers of artificial neurons, which transfer the characteristics to the predicted network output [13].

Literature Survey:

1 Hideaki Yanagisawa, Takuro Yamashita, Hiroshi Watanabe proposed A Research on Object Identification Methods from Manga Pictures. In this Paper they investigated the usefulness of object proposals for manga object identification in this work. The experimental results reveal that Selective Search is successful for things with clear borders while RPN is useful for objects with uncertain bounds. Furthermore, because it is difficult for SSD to detect manga things in the entire image, the procedure of partitioning the image into tiny areas is required.

2 Reagan L. Galvez, Argel A. Bandala, Elmer P. Dadios, Ryan Rhay P. Vicerra, ose Martin Z. Maningo. Detecting Objects using Convolutional Neural Networks proposed The object identification capabilities of the two cutting-edge CNN models was effectively shown. It demonstrates that the SSD with MobileNetV1 has high speed detection but low accuracy when compared to Faster-RCNN, which has low speed but higher accuracy. According to the findings of the trials, there is a trade-off between accuracy and speed. Use SSD with MobileNetV1 if

we desire quick detecting capabilities, especially in real-time applications. Faster- RCNN with InceptionV2 is recommended for high detection accuracy. The two models will be used in future studies as the vision system of a bomb disposal robot to detect improvised explosive devices.

Problem Identification

present day Real-time object identification helps with cause estimation, vehicle detection, traffic control, CCTV surveillance, and other applications. As a result, day-to-day growth in Technology is required for the capacity to provide vision in a computer. As a result, low-cost technology and learning tools are required. The current computer vision approaches are more expensive than prior techniques due to the development of bigger and deeper networks to achieve improved accuracy. Methods that are totally computer-based need a large amount of GPU power and aren't necessarily real time, making them unsuitable for everyday applications. Hence, in this project, we propose a model called "Object detection," which is a CPU-based YOLO that uses OpenCV to execute the YOLO

method on non-GPU computers and can be utilised on standard systems without the need for any additional added features.

Methodology

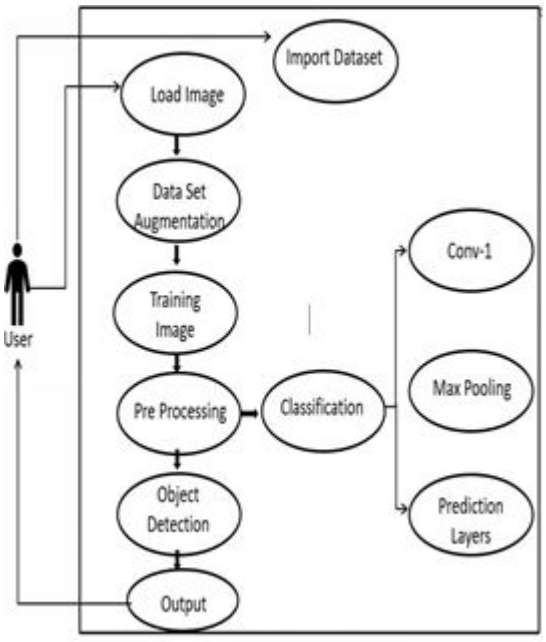
Step 1: Gathering data and creating a dataset This requires pulling photos from the Kaggle database, analysing them, and then extracting features.

Step 2: Creating a Convetional Neural Network for different kind of Images and videos. For the purpose of detecting objects from different images and videos using conventional neural network model is created in this stage.

Step 3: Training and testing with datasets

The conventional neural network model is tested and trained on the datasets to detect or recognize objects in image or video

Step 4: Deployment and analysis using use case diagram



Implementation

To follow along we will need the following:

1. Python 3.10 or above installed on your device

2. A Dataset available in the form of images and videos

The categorization of images is a frequent application for convolutional neural networks (CNNs), a form of neural network. They employ a number of convolutional and pooling layers in order to identify patterns in pictures. The fully connected layer predicts after the pooling layer has downscaled the image and the convolutional layer has applied filters to extract features from the input image. The optimal filters are taught to the network via backpropagation and gradient descent.

Python implementation using CNN

In Python, we can implement a CNN using Numpy and then we Load the preprocessed data and later we define cnn model using Numpy Define the functions for convolution, pooling and activation. The modules imported from Python is OpenCV. OpenCV: Face recognition, machine learning, artificial intelligence, and other fields use OpenCV, a Python open-source package, for computer vision. The term "computer vision," abbreviated as "CV" in OpenCV, refers to a branch of research that enables computers to comprehend the content of digital images like pictures and movies. In order to comprehend the content of the images, computer vision is used. It pulls out the textual descriptions, three-dimensional models, item descriptions, and other information from the images.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Results & Conclusion

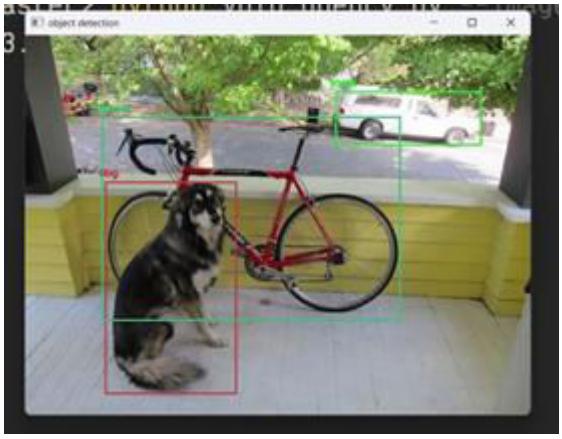


Figure1: In this image we are detecting the dog, cycle and car as objects.

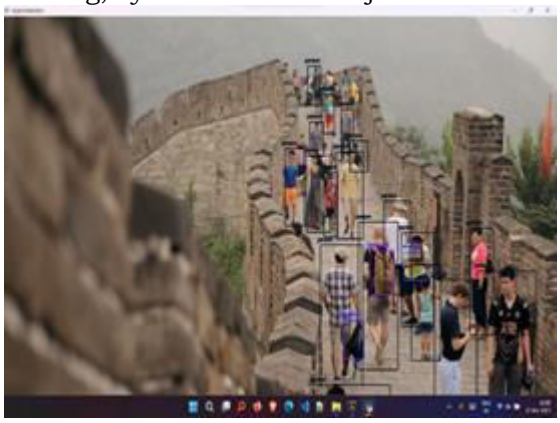


Figure2: In this video we are detecting different persons and their belongings as objects.

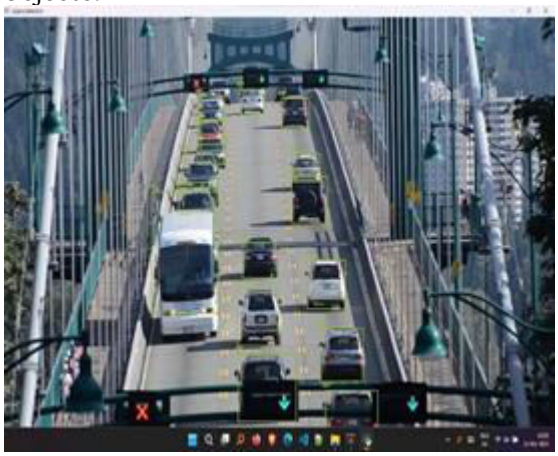


Figure3: In this video we are detecting different vechiles and recogning them as car, truck, bus, etc,.

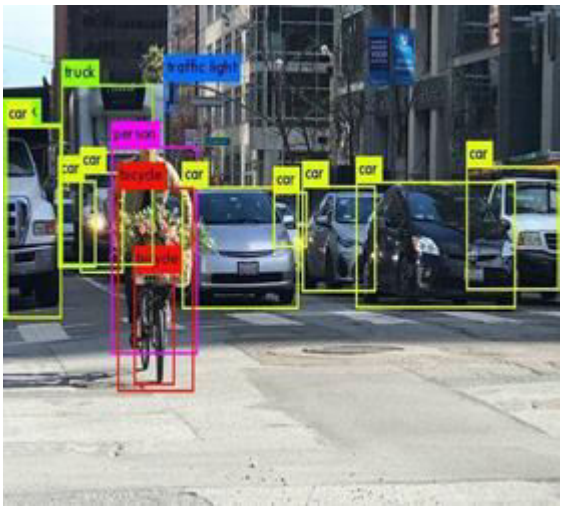


Figure4: In this video we are detecting different kind of objects at a traffic signal.

Conclusion

We primarily concentrated on developing a straightforward, user-friendly object identification model for movies in this research. Yolo and OpenCV are the only two components of this extremely lightweight model. tracking down the things in the movies. For real-time object identification on non-computers, CPU Based YOLO is advantageous. Before using the YOLO technique correctly and successfully completing the task, we first chose the best framework for doing the object detection task on the CPU. In order to identify items in photos, we have created an algorithm employing deep learning neural networks. In order to detect objects with high accuracy and in real time, the research employs the single shot multi-box detector (SSMBD) method in combination with Faster CNN. With both static photos and moving pictures, our algorithm performs well. The accuracy of the suggested model is higher than 70%. This model extracts

feature data from the image using convolutional neural networks, which is then converted into a class label via feature mapping. To address the variation in aspect ratio, our method employs several filters with unique default boxes and also makes use of multi-scale feature maps.

Limitations & Future Work

Object classification and object location are the two main obstacles in object detection. To address these problems, researchers are employing a multi-task loss function. This multi-task loss function aids in producing consequences for both localization mistakes and misclassification errors. Some items are just partially visible, making their detection challenging. Objects that take up more space are simpler to grasp than those with smaller spaces, which make it easier to perceive them. It is difficult to find tiny objects in any image or video.

In future, we will perform the tasks by training the machine with custom dataset.

References

[1] V. Gajjar, A. Gurnani and Y. Khandhediya, The article "Human Identification and Tracking for Video Surveillance: A Cognitive Science Approach" was published in the 2017 IEEE International Conference on Computer Vision Workshops.

[2] Srinivas S, Sarvadevabhatla RK, Mopuri RK, Prabhu N, Kruthiventi SSS, Venkatesh Babu R. An introduction to deep convolutional neural nets for computer vision. In: Deep Learning for Medical Image Analysis. Academic Press; 2017. pp. 25-52

[3] Noh, H., Hong, S. and Han, B. (2015) The Learning Deconvolution Network for Semantic Segmentation. 1520–1528 are included in the proceedings of the IEEE International Conference on Computer Vision, held in Santiago from December 7– 13,2015.

[4] C.-C. Kao, P. Sen, T.-Y. Lee and M.-Y. Liu, "Localization-Aware Active Learning ObjectDetection,"arXiv:1801.05124v1, Ja nuary2018.

[5] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y. and Berg, A.C. (2016) SSD: Single Shot Multibox Detector. European Conference on Computer Vision, Amsterdam, 8-16 October 2016, 21-37.

[6] Pan WD, Dong Y, Wu D. Classification of malaria-infected cells using deep convolutional neural networks. In: Machine Learning: Advanced Techniques and Emerging Applications. 2018. p. 159

[7] Tan, M. and Le, Q. (2019) EfficientNet: Convolutional neural networks: Rethinking model scaling. Long Beach, CA, 10-15 June 2019, 6105–61114, International Conference on Machine Learning.

[8] H. Xu, X. Lv, X. Wang, Z. Ren and R. Chellappa, "Deep Regionlets for Object Detection,"arXiv:1712.02408v1, December 2017.

[9] Girshick R. Fast r-cnn. In: The 2015 IEEE International Conference on

Computer Vision Proceedings, pp. 1440–1448

[10] Ren S, He K, Girshick R, Sun J. Faster r-cnn: Using region proposal networks, we are moving towards real-time object detection. Pages. 91–99 in Advances in Neural Information Processing Systems, 2015.

[11] Zeiler, M.D. and Fergus, R. (2014) Convolutional networks: Visualizing and Understanding. Zurich, Switzerland, September 6–12, 818–833. European Conference on Computer Vision.

[12] Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: Unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016. pp. 779-788

13] G. Huang, Z. Liu, L. Van der Maaten and K. Q. Weinberger, "Densely Connected Convolutional Networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017